U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
NATIONAL METEOROLOGICAL CENTER

OFFICE NOTE 369

ONE YEAR OF OPERATIONAL COMPREHENSIVE HYDROSTATIC QUALITY CONTROL AT THE NATIONAL METEOROLOGICAL CENTER

LEV S. GANDIN
WILLIAM G. COLLINS

DEVELOPMENT DIVISION

MAY 1990

THIS IS AN UNREVIEWED MANUSCRIPT, PRIMARILY INTENDED FOR INFORMAL EXCHANGE OF INFORMATION AMONG NMC STAFF MEMBERS.

1. Introduction

The Comprehensive Hydrostatic Quality Control (CHQC) of rawinsonde data on height and temperature at mandatory isobaric surfaces has been designed and implemented into routine operational practice at the National Meteorological Center (NMC). The principles of the CHQC approach and its methodology, including the Decision Making Algorithm, have been described in detail elsewhere (Collins and Gandin, 1990, hereafter refered to as CG). The purpose of this paper is to present the statistics of the CHQC performance based on results of its operational use at NMC for one year, January to December 1989.

The aim of collecting and analysing these statistics was twofold. First, such statistics provide a good base, maybe even the only one, for any improvement of the CHQC algorithms. In fact, an improved CHQC version replaced the previous one in July 1989. Despite substantial changes in the CQHC procedures, the results did not show very much difference, which simply means that the previous version was not bad. The overall statistics of the CHQC performance may be therefore based on its results for the whole year. Some more subtle effects have been, however, analysed using only the improved CHQC version.

We have to stress in this respect that every conclusion about a desirable change in a quality control (QC) procedure should be based on statistics of its application to a large amount of data, not to a single case. This is particularly so because an overwhelming majority of data do not contain, of course, rough errors detectable by the QC, so that in order to investigate the QC reaction it is necessary to apply it to very many cases. This is true for any QC method, not only for the CHQC.

The second aim was quite different and perhaps even more important: to gain information on the present status of the data quality, on the frequency and geographical distribution of rough errors detectable by the CHQC and on their origin. In other words, the aim was to perform the data quality monitoring based on the CHQC performance statistics.

The problem of data quality monitoring has attracted much attention in recent years both at NMC and elsewhere, and substantial amount of work in this direction has been done particularly at the European Centre for Medium Range Weather Forecasts (Hollingsworth et al, 1986, Bottger, Radford, and Soderman, 1987). These works were, however, directed, almost exclusively, towards investigation of systematic errors. Such errors are, of course, very important despite being comparatively small, because they persist in time. As to the rough errors, very little had been known about the present situation with NMC before the CHQC began to be applied on a regular basis at NMC.

Having information on the statistics of "hydrostatic" errors (i.e., errors detectable by the CHQC), it is possible to propose some measures desirable in order to improve the situation, to diminish the number of such errors. Many proposals of this kind were made based on the CHQC monitoring statistics. Some of the proposed actions have already been undertaken, some are underway, but most of the proposals are still not considered. Many such proposals will be discussed in this article.

Although each rough hydrostatic error is due to some definite cause and is, in this sense, not random, the whole set of rough errors behaves like that of random errors. Therefore, reliable statistics on their behavior may be obtained only by using samples that are large enough. In other words, substantial averaging both in time and in space is needed in order to obtain statistics we may believe in. This is particularly so because rough hydrostatic errors, as well as rough errors in general, occur comparatively seldom.

Table 1 LARGE REGIONS

Description	Notation	WMO Blocks	Mean number of reporting stations 00	12
Western Europe	W Eu	1-4,6-8,10,16	62	71
Eastern Europe	E Eu	9,11-13,15,17	23	24
USSR	USSR	20-38	167	170
Western Asia	W As	40-41	18	14
India, Ceylon	Ind	42-43	25	23
Mongolia	Mong	44	7	7
Taiwan, Korea, Japan	Jap	45-47	30	30
Indochina, Malaysia	Indo	48	10	8
China	Chin	50-59	119	122
Northern & Central Africa	Af	60-65,67	14	24
South Africa	S Af	68	11	12
USA	USA	70,72,74	82	90
Canada	Can	71	30	32
Central America	C Am	76,78	8	20
South America	S Am	80-87	6	21
Antarctica	Ant	89	11	6
Pacific Islands,Indonesia	P Is	91,96-98	32	11
New Zealand, Australia	Aust	93-94	36	13

We did not consider therefore time averaging intervals less than one month, and we also tried to estimate the representativeness of such statistics by comparing them for different months. As to spatial averaging, we first tried to perform it by so-called WMO blocks which divide the globe in slightly less than 100 areas. It became clear very soon, however, that this was not suitable for our purpose: random variability of such statistics from month to month was too large for almost all WMO blocks. We decided therefore to average over larger areas, each consisting, as a rule, of several blocks and chosen in such a way that there should be no substantial systematic variability of the CHQC statistics within any of the areas. These areas are called large regions, and Table 1 presents WMO blocks for each large region, as well as the average number of reports per observation time (00 UTC and 12 UTC) received at NMC from each large region.

Most of the statistics presented below were obtained entirely automatically, by the CHQC Performance Statistics Code, which also produces detailed summaries for each month. Investigation of several effects required, however, some subjective judgement and has been therefore performed manually. As a rule, only some part of all available data was used for every such evaluation, but we tried in each case to estimate the degree of confidence in these results.

Unlike other quality control methods, the CHQC tries to discover the cause of each suspected error. As is well known, all rough errors in rawinsonde information, as well as in any other data, may be divided into 3 categories: (1) observation (or, better to say, measurement) errors, when a sensor gets out of order or a measurement result is distorted before it enters the processing at the station (or elsewhere); (2) computation errors, introduced in the course of this processing; and (3) communication errors, originating when data are put into, follow along, and are taken from communication lines.

As long as only two of three parameters - pressure, temperature, and height - are measured directly, and the hydrostatic equation is used to compute the remaining parameter, usually the height, while processing the data, the CHQC is incapable of detecting observation errors; it simply does not react to them. As to computation errors in the course of this processing, each error of this kind results in a wrong value of one thickness and, consequently, in wrong heights of all mandatory surfaces above. CHQC detects such errors, but it would be too risky to make automatic corrections of many heights based on only one large residual, and the Decision Making Algorithm (DMA) only displays the suspected error in the thickness computation. All other errors detectable by the CHQC are communication related.

Attempting to discover the cause of each probable error, the DMA assigns one or another Error Type to each suspected error or errors. These types,.listed in Table 2, were discussed in detail in CG. As mentioned there, the CHQC DMA performs automatically only so-called confident corrections (those of error types 1, 2, and 7 - 10) while for errors of all other types, it only displays diagnoses without changing any data. (Moreover, type 7 - 10 errors were not corrected before July, when the previous CHQC version was operational).

Table 2 TYPES OF HYDROSTATIC ERRORS

1. Confident height correction.
2. Confident temperature correction.
3. Both temperature and height of the same level suspected.
4. Height or temperature of the lowest level (or both) wrong, or an error in the lowest layer thickness computation.
5. Height or temperature of the highest level (or both) wrong.
6. Thickness computation error suspected.
7. Confident corrections of two adjacent height errors.
8. Confident corrections of two adjacent temperature errors.
9. Confident corrections of two adjacent errors: height below and temperature above.
10. Confident corrections of two adjacent errors: temperature below and height above.
11. Suspected height error, not large enough to be confidently corrected by the hydrostatic check alone.
12. Temperature correction not made because this would produce an unstable vertical temperature profile.
13. Hydrostatically detected data hole in the upper part of the Part A, with 100 mb absent and at least one part C level present.
14. Hydrostatically detected data hole, other than Type 13.
22. Suspected temperature error, not large enough to be confidently corrected by the hydrostatic check alone.
99. Type 8 or 9 or 10 correction not made because this would produce an unstable vertical temperature profile.

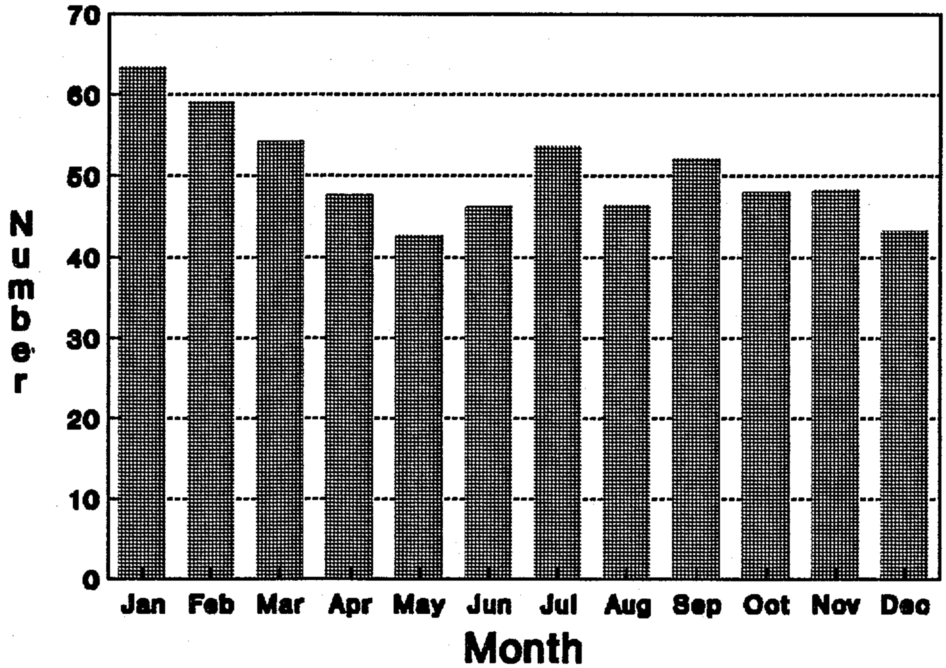

Fig.1. Monthly mean numbers of suspected hydrostatic errors per observation time.

2. General statistics of hydrostatic errors.

2.1 Overall numbers and geographical distribution. Averaged over a month, 40 to 65 suspected hydrostatic errors occured for each main observation time (00Z UTC, 12 UTC) which amounts to 6.5 - 8.0 percent of all rawinsonde reports received at NMC (Fig.1). These figures vary from month to month, but variations are comparatively small and rather irregular. It would be, of course, wrong to say that about 7 % of all received data are erroneous. Because each report contains many values and only a small part of them, most often only one, is wrong, the real number is much smaller. Nevertheless, the numbers in Fig. 1 show that hydrostatic errors still occur comparatively often despite all achievements in automation of observations, processing, and communication.

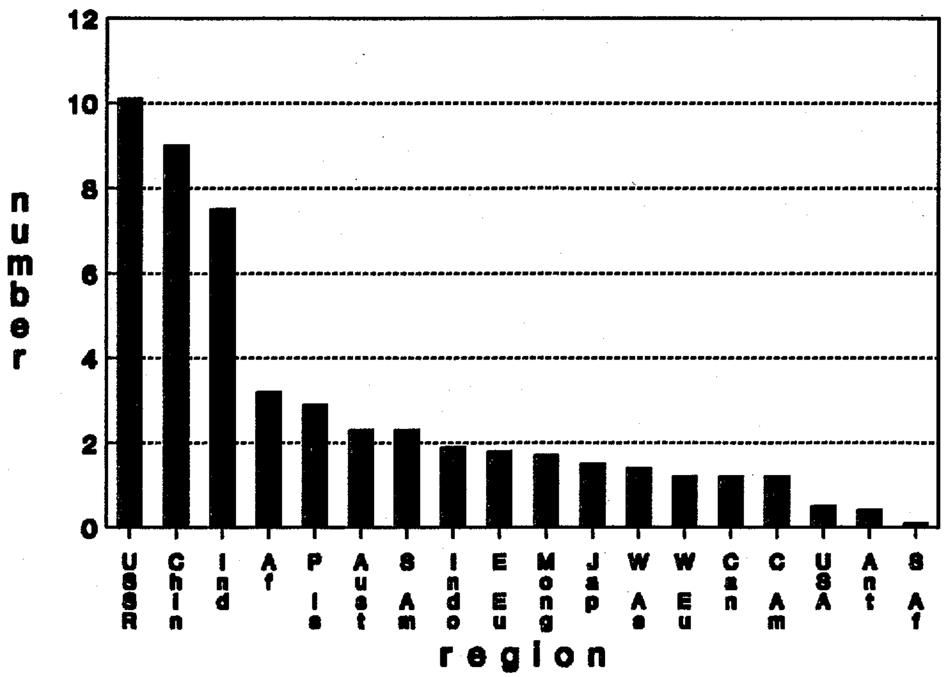

Fig.2. Annual mean numbers of suspected hydrostatic errors per observation time by large regions.

The explanation is very simple: as may be seen from Fig.2, an overwhelming majority of hydrostatic errors originated in countries where the communication processes are still not computerized, or computerized incompletely, so that many operations are performed manually. For example, the Soviet Union, India, and China produce about a half of all errors, while the error numbers over United States and over some West-European countries, where both rawinsonde data processing and communication lines perform completely automatically, are very small.

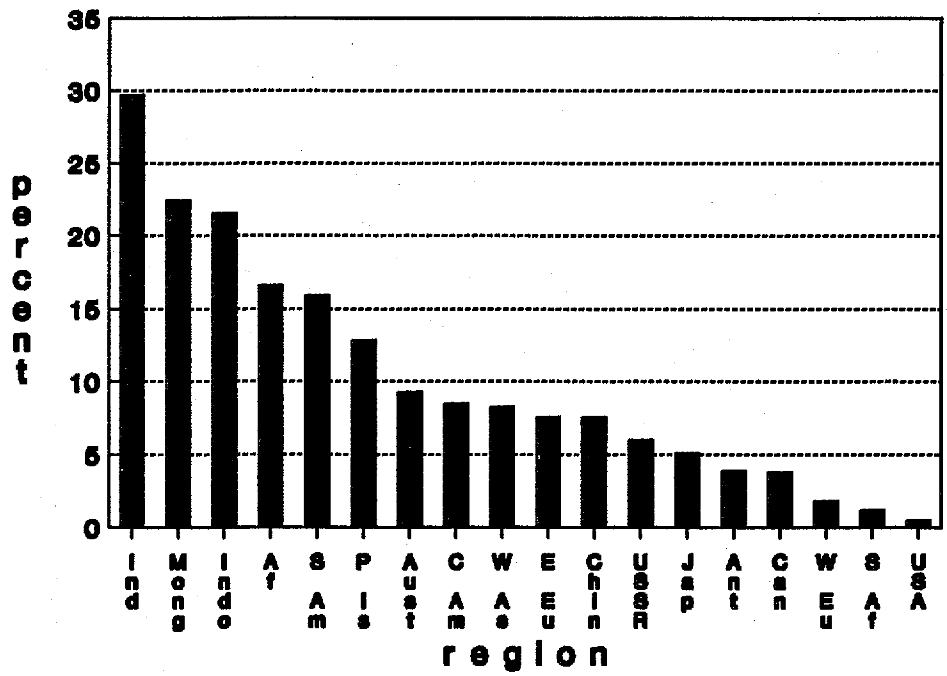

Fig.3. Annual mean percent of error-containing reports by large regions.

Comparison of absolute numbers of errors for different regions is, however, somewhat misleading, because they depend on the number of reporting stations, and these numbers are quite different for different large regions. In that sense, information on the percentage of error-containing reports, presented in Fig.3, gives a more realistic view. As one can see, the relative frequencies of errors over USSR and China are not so large, as would seem from Fig.2, while these frequencies over Indochina and Mongolia are substantially larger than it might seem. The only exception in this respect is India: its error numbers are very large from both absolute and relative points of view.

It is necessary to take into account that rough errors in areas with a poor data coverage are much more dangerous than those over regions with a dense rawinsonde network. First, an isolated rough error in a data-rich region influences the results of objective analysis and, thus, the forecast to much lesser extent as compared with such influence in a data-poor domain. Secondly, one may reject a correctable erroneous datum over a region with a dense network, or even blacklist a station for some period without causing any substantial harm (although it is, of course, better not to lose information if it can be corrected). The situation for a data-poor area is quite different. Rejection of a single report in such an area would often lead to absence of any data over a large domain. Every effort should be made therefore to correct erroneous data over sparse network regions or, at least, to reject as few data as possible in such regions.

Along with the geographical variability of the error frequency from one large region to another, there exists a pronounced variability from one station to another. While, in global average, one among fifteen reports or so contains a suspected rough hydrostatic error, there are some stations with much higher error frequency. Stations with more than 100 suspected hydrostatic errors for the period from May to December 1989 (which means one error per, at most, five reports) are listed in Table 3. This inhomogeneity can hardly be ascribed to anything but a lack of training of personnel at some stations. A feedback with countries having such stations may therefore be useful for improvement of the situation. It would be highly desirable to have such information disseminated regularly, preferably under the WMO guidance.

Table 3 STATIONS WITH MORE THAN 100 HYDROSTATIC ERRORS
May - December 1989

No	Index	Name	Country	N
1	46747	Tungkong	Taiwan	207
2	43371	Trivandrum	India	202
3	43346	Kairakal	India	177
4	44354	Sainshand	Mongolia	176
5	43369	Minicou	India	174
6	43041	Jagdalpur	India	160
7	43185	Machilipatnam	India	157
8	54337	Chinchow	China	156
9	67197	Fort Dauphin	Malagasy Rep.	154
10	43128	Begumpet	India	151
11	60680	Tamanrasset	Algeria	142
12	42101	Patiala	Bangladesh	136
13	42410	Gauhati	India	133
14	43285	Panambur	India	130
15	42971	Bhubakeshvar	India	128
16	42867	Nagpur Sonegaon	India	126
17	43295	Bangalore	India	123
18	31538	Sutur	USSR	122
19	44373	Dalanzadgad	Mongolia	120
20	41780	Karachi	Pakistan	117
21-22	30673	Mogacha	USSR	116
21-22	82193	Belem	Brazil	116
23-26	38750	Gasan-Kuli	USSR	114
23-26	42369	Luknow	India	114
23-26	44212	Ulan Gom	Mongolia	114
23-26	91643	Funafuti	Kiribati	114
27	48327	Chiang Mai	Thailand	112
28	55299	Heiho	China	110
29-31	15730	Kardjali	Bulgaria	108
29-31	24688	Oimyakon	USSR	108
29-31	42667	Bhopal	India	108
32	20107	Barentsburg	USSR	107
33	96471	Kota Kinabalu	Brunei	104
34	42647	Ahmedabad	India	103
35-36	43279	Madras	India	101
35-36	56137	Changtu	China	101

2.2. Confident corrections.

Some statistics concerning hydrostatic errors of various types are presented in Table 4. One can see from the Table that about 50 percent of all suspected errors were automatically corrected by the CHQC DMA. This percentage slightly increased in time due to some improvements of the DMA.

Table 4. Monthly Mean Numbers of Suspected Errors of Various Types

type	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	16.3	15.8	15.1	13.6	13.2	13.4	14.5	12.9	14.6	13.7	15.0	14.7
2	11.9	12.0	12.1	11.0	9.9	10.8	12.8	12.8	12.2	11.7	12.9	10.7
1+2	28.2	27.8	27.1	24.6	23.1	24.2	27.3	25.6	26.8	25.4	28.0	25.4
other	35.1	31.2	27.0	23.0	19.5	22.0	26.2	20.6	25.3	22.6	20.2	17.8
3	4.8	5.2	4.3	3.9	3.9	3.3	1.6	1.3	1.7	1.7	0.9	1.0
4	6.3	6.3	5.5	4.3	3.9	3.3	3.8	2.1	1.2	2.2	1.7	1.3
5	8.2	5.3	4.1	3.8	3.7	3.9	2.9	2.1	2.7	2.7	2.2	1.9
4+5	14.5	10.8	8.4	7.7	7.0	7.7	5.0	3.4	4.9	4.4	3.5	3.5
6	7.7	7.1	6.7	5.3	3.8	5.4	3.4	3.1	3.0	2.6	2.8	1.7
7-10	2.2	2.3	1.2	1.4	1.2	1.6	2.0	1.6	2.0	1.6	1.9	1.6
11	4.1	3.9	4.0	3.1	2.8	2.9	3.7	2.7	3.1	3.0	2.8	3.1
13							6.6	6.0	7.8	6.9	5.7	5.0
14							0.6	0.5	0.6	0.6	0.6	0.6
22							0.7	0.8	0.8	0.6	0.9	0.4

The numbers of confident height corrections (Type 1) regularly exceeded, at least slightly, those of confident temperature corrections (Type 2), which may be explained by the fact that temperatures look more familiar to those communicating the reports, and the probability for them to make a rough error in temperature is less than in height of an isobaric surface. As to other confident corrections, those of Types 7 to 10, they occur, of course, much more seldom than Type 1 or 2 corrections.

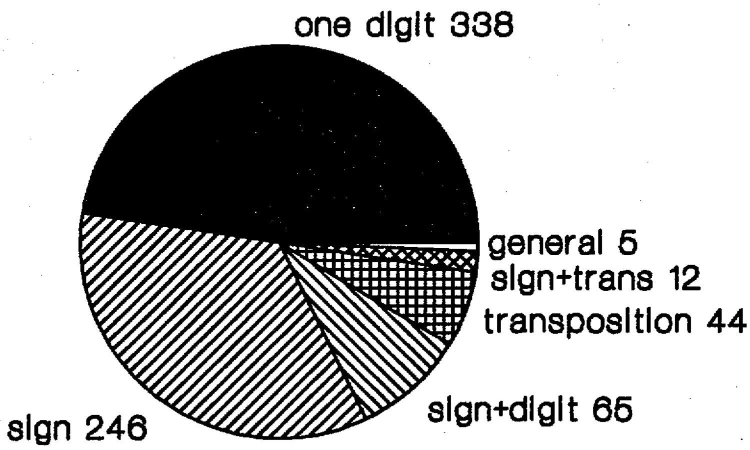

Fig.4. Monthly mean numbers of various Type 2 corrections, November and December 1989.

As discussed in some detail in CG, the CHQC Decision Making Algorithm tries, after having detected each error, to find a "simple" correction, that is, a correction changing only one digit in reported value, or only sign (for temperature), or both, or resulting only in transposition of digits (maybe in combination with the sign change). Some statistics of various kinds of confident temperature corrections, collected for two months, November and December 1989, are presented in Fig.4. They show, first of all, that almost all confident temperature corrections turn out to be simple ones: only about 2% of these corrections are general type ones. One digit temperature corrections and sign corrections take place most often, forming, resp., about 46% and 37% of the whole number of confident temperature corrections. Transposition correction and "combined" ones (digit plus sign, transposition plus sign) happen much more seldom, but still not so seldom as those of general type.

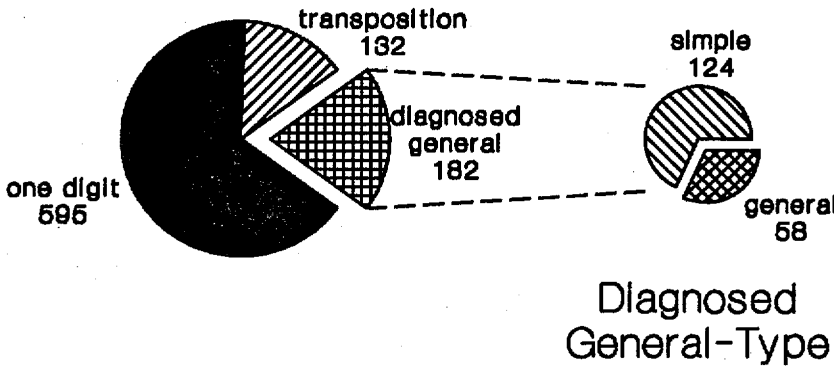

Fig.5. Monthly mean numbers of various Type 1 corrections, November and December 1989.

It may seem that the distribution is quite different for confident height corrections (Fig.5): as many as about 20% of them are diagnosed by the DMA as general-type corrections. A more thorough analysis shows, however, that the majority of them are actually simple corrections, but different from one-digit or transposition ones. For example, in some cases, the correction results in exclusion of a digit contained in the reported height value, or in insertion of an additional digit. There are many other kinds of such "general but simple" height correction. The frequency of them is also shown in the figure as well as that of remaining general-type corrections. The latter is small, though still higher than it is for temperature corrections.

The conclusion is that an overwhelming majority of both Type 1 and Type 2 errors are simple errors, and this proves again that the errors are made during manual operations. Even if we had no independent information that the communication procedures in many countries are still not completely computerized, we could make this conclusion from presented statistics, just because computers have nothing to do with decimal digits.

We have also computed some statistics of multiple confidently corrected errors, that is, of reports containing two or more such errors each. We already had an impression that such reports happen much more often than one could expect, and this led us to the decision to make, in the CHQC DMA, special provisions for cases of two errors at neighboring levels, i.e., for those of Types 7-10. Even more often, there occur two or more isolated errors, that is, errors not at neighboring levels, in the same report, which do not need special DMA provisions.

The averaged monthly number of confident corrections is about 1560, or about 3.5% of all reports. If these errors were made independently, then the monthly number of reports with two confidently detected errors would be about 55. The actual number, averaged over the same months, November and December 1989, was about 130, or more than twice as large.

In order to explain this effect, one may present the probability of a confidently correctable error, p, as a product of two probabilities: probability p_1 that the station has an error producing personnel, and probability p_2 that they will produce an error. In that case the probability of a report with two errors is not $p^2=(p_1)^2(p_2)^2$, but $p_1(p_2)^2$. Using the estimates above, we obtain $p_1=0.42$; $p_2=0.087$. This means that about 40% of personnel involved are capable of producing confidently correctable errors, and each of them does so in about 9% of reports. We are, certainly, not insisting on these very rough estimates, they just illustrate what may be happening. At the same time, the monthly average number of reports with triple corrections should be, according to these estimates, about 11, which is not far from the actual number.

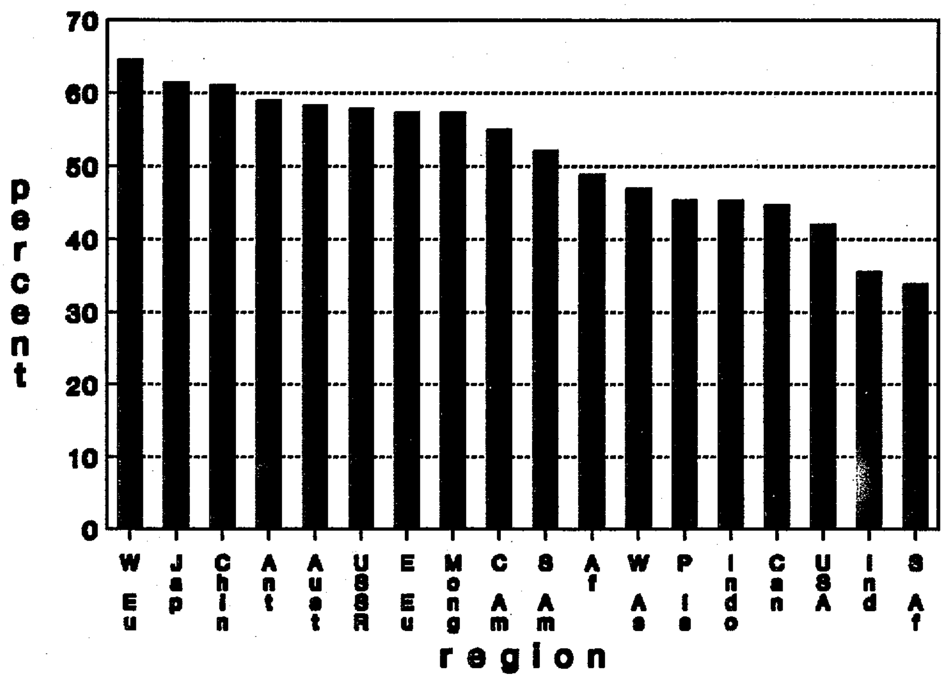

Fig.6. Annual mean percent of confident corrections by large regions.

There also exists, as may be seen from Fig.6, a persistent geographic non-homogeneity in the percentage of confidently correctable errors. It usually exceeds 60% over, say, Western Europe and Australia, which indicates that most hydrostatic errors over these regions are introduced in the course of communicating the data. At the same time, the percentage is much smaller, usually less than 40%, over India, and this is another proof that many hydrostatic errors are caused by improperly trained personnel.

2.3. Non - confident corrections.

As long as hydrostatic checks in the CHQC are not being accompanied by other, statistical, checks, there exist many types of suspected errors for which the DMA, as a rule, only proposes some corrections but does not introduce them. Outputs with information on such non-confidently correctable errors are being manually analysed by specialists at the NMC Meteorological Operations Division in the course of their real-time actions, and the final decision for every such datum is made by them. We also analysed, in a quasi-operational monitoring regime, these data (together with those for confident corrections), in order to collect sufficient statistics and to improve the DMA. What we, as well as MOD specialists, actually did, is something like a subjective vertical consistency check of full values (i. e., not of their deviations from any background profiles) for both height and temperature.

If a suspected communication error satisfies both existence and magnitude conditions, as described in CG, but is still not large enough to be confidently corrected, then it is assigned Type 11 for height or Type 22 for temperature (see Table 2 on p. 4). The present upper limits for Type 11 error absolute values are 29 m for levels 1000, 850, and 700 HPa (where numbers of meters are known from reports), and 80 m above (where only decameters are known). As to Type 22, the upper limit everywhere is 9.9^{o}; in other words, if a temperature correction is 10^{o} or more, it belongs to Type 2, not 22, and is automatically performed, not only displayed.

As it may be seen from Table 4 (p. 10), errors of these two types occur comparatively often, and we tried to recognize what has actually happened in every such case. In some of these cases, it was more or less clear that even a small correction should be accepted, like the sign correction in temperature -4.3^{o}. In other cases, however, even larger corrections looked questionable. The general conclusion is that not only a vertical statistical check, which we tried to simulate subjectively, but also horizontal and/or temporal checks are to be added in order to resolve such cases.

What should be done if the DMA diagnoses errors at two adjacent mandatory levels (an error of any of Types 7 - 10), and one of the two errors is small so that, if present alone, it would belong to Type 11 or 22? Extensive testing showed that both corrections, including the small one, have to be done in such cases. As opposite to that, if one of the two corrections would result in a strong superadiabatic lapse rate or in an excessive curvature of the temperature profile (Type 12 error), then neither of the corrections should be made. More generally, if the DMA diagnoses a Type 12 error, this is practically always an indication that either another correction, or no correction at all should be made. The same is true, as a rule, with Type 3 errors, when the DMA decides that *both* height and temperature at the same level are wrong and proposes corrections to make both residuals equal to zero. Such errors in both parameters do happen, but very seldom. More often, the Type 3 diagnosis occurs when there are several errors in such combination that the DMA cannot confidently correct all of them. Sometimes, such cases may be resolved by specialists, but that usually requires much effort and even some imagination.

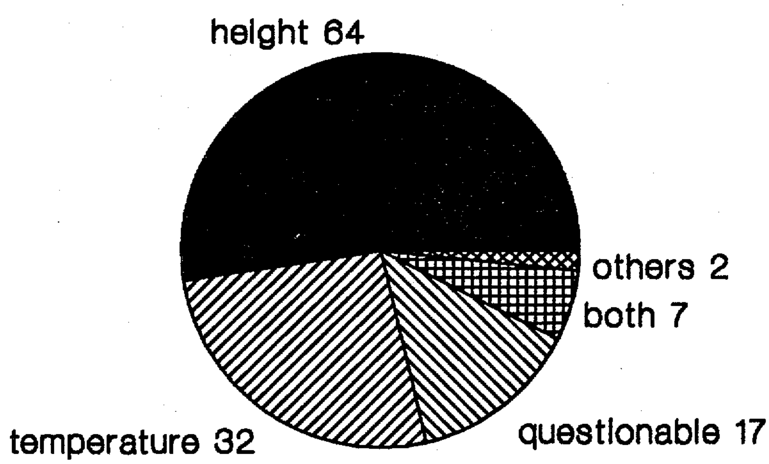

Fig.7. Monthly mean numbers of various Type 5 errors, November and December 1989.

In contrast to Type 3 errors, Type 4 and 5 errors can usually be easily and quickly treated by specialists. Type 5 is a suspected error in height *or* temperature of the uppermost level (among those reported). The DMA presents possible corrections of either of them, and the specialist has only to decide which of them, if any, to chose. Fig.7 shows that about 80% of Type 5 errors may be confidently corrected this way under operational conditions. In addition to this, both height and temperature were found erroneous in about 5% of Type 5 errors, and both were therefore rejected, in such cases by the specialist. There were thus only 15% of Type 5 cases that could not be decided upon univaluedly.

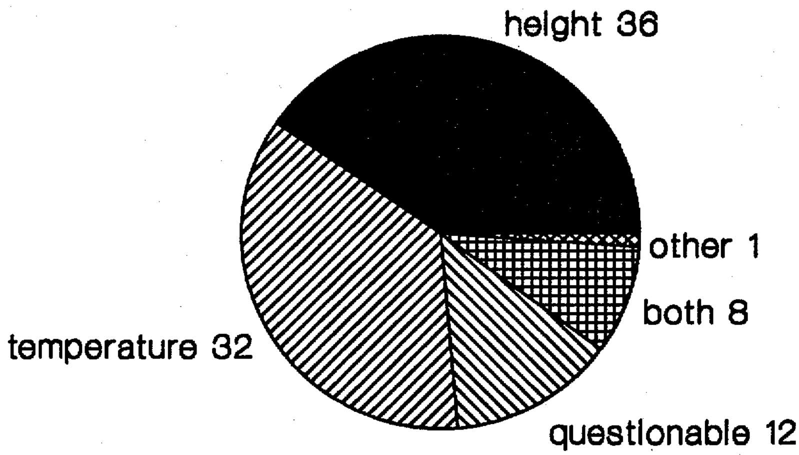

Fig.8. Monthly mean numbers of various Type 4 errors, November and December 1898.

A similar situation takes place in cases with Type 4 errors, when the DMA proposes alternative corrections for height or temperature of the lowest level (Fig.8). For these errors, some additional information may result from the baseline check (see section 3.4), which computes possible corrections of 1000 mb or 850 mb height (along with those of surface pressure and station elevation) in every case of suspected baseline error. The statistics of Type 4 error corrections does not practically differ from that of Type 5.

It should be mentioned that a Type 4 error may be also caused not by a communication error, but by a computational error in the thickness of the lowest layer, so that all the heights, except the lowest one, are to be corrected (this may also be the cause of a Type 5 error, but that leads only to an error in the height of the upper level, just like a communication error in this height). This may be considered as a particular case of Type 6 error - a suspected error in thickness computation.

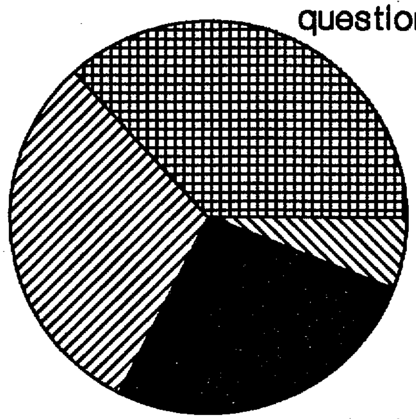

Fig.9. Monthly mean numbers of various Type 6 errors, November and December 1989.

A Type 6 error occurs when a hydrostatic residual for one layer is large while those for neighboring layers are small. If the large residual exceeds, say, 200 m, then one may be almost sure that the thickness was computed (or written down) wrongly in the course of the mandatory level height computations. Even in those cases, however, it would be risky to correct several heights based on only one residual. There also exists another possible origin of suspected Type 6 errors, that is the non-linearity of the temperature profile (with respect to the logarithm of pressure) within the layer. If the residual is not very large, it may happen to be caused by the non-linearity only, so that no correction is needed at all. We tried to discover, in the course of our quasi-operational monitoring, what has actually happened in each case with a diagnosed Type 6 error. Statistics based on these attempts (Fig.9) show that about a third of these errors are definitely errors in thickness computation, a third are, most probably, caused by the temperature profile non-linearity, and for the remaining third, it is difficult to reach a definite conclusion.

It is necessary to stress again, in this respect, that by the introduction of even simplified vertical and, particularly, horizontal and/or temporal interpolation checks in addition to the hydrostatic one, it will be possible to make univalued decisions in an overwhelming majority of Type 6 errors, as well as those of Types 4 and 5. A majority of Type 3, 11 and 22 errors will be also treated entirely automatically. Even more important, only by this way will it be possible to detect measurement errors. The design of a comprehensive quality control algorithm containing both horizontal and vertical interpolation checks is now underway, and we hope to complete the design in several months.

3.Examples of problems discovered with the CHQC.

3.1. CHQC as an instrument for data monitoring.

At the very beginning of the CHQC functioning, quite unexpectedly for ourselves, we came across some shortcomings in data entering the Data Assimilation Systems at NMC. Some of these deficiencies existed for several decades but were never recognised before. We attracted the attention of specialists involved to such undesirable problems and, as a result, some of these problems no longer exist. There also were many cases when one or another distortion of data suddenly occured, and again, it was the quasi-operational monitoring of the CHQC performance that made it possible to discover such events comparatively soon and, finally, to get rid of them. All that proves that the CHQC, though being directed towards exclusion of rough errors, may be efficiently used as a means for data quality monitoring. This is true, of course, for any quality control procedure, but the fact that the CHQC tries to find out, in each case, the cause of suspected error makes it particularly useful as a data monitoring instrument.

We are not going, of course, to describe here in detail every particular problem of this kind we have come across so far. What follows should be considered rather as illustrations of the CHQC application for data monitoring, with particular emphasis on actions desirable in order to make the occurence of such detrimental problems in the future as rare as possible.

3.2. Data holes.

A strange effect which we called "data holes" demonstrated itself from the very beginning of the CHQC testing: data on both temperature and height were missing for several mandatory surfaces in a row. Most often, this occured with a sequence of upper levels in the part A of the rawinsonde report (ending at 100 HPa), while there existed some data reported in the part C (Type 13 error). Data holes also occured which were not connected with the division of the mandatory level information between parts A and C (Type 14), but they were much more seldom.

It was quite easy to discover the origin of this phenomenon: the deficiency of the computer program for decoding rawinsonde messages (the decoder). Each time the decoder comes across some code violation in a report, it ceases its decoding not only of anything else at the level where the violation occurred (which may be also avoided quite often), but of the whole remaining part of the report! Fortunately, part C is transmitted not immediately after part A, and the decoder deals with part C independently of what it has done with part A. This explains why it happens so often that several levels, up to 100 HPa, are missing in a decoded message, while some levels above, usually beginning at 70 HPa, do exist.

The decoder has been in use for more than a decade, and one may be sure that there always were many data holes. Moreover, when this phenomenon was discovered, it has been shown that data holes happen much more often than this may seem from the CHQC results: in cases when there is no information above a hole, or when the hole causes no large hydrostatic residual, the CHQC is unable to discover the hole.

The decoding of rawinsonde reports was, and continues to be, performed outside NMC, by a program written in an assembler language and not well documented. Table 4 shows that the frequency of Type 13 errors remained high and practically constant from month to month in 1989. It should be added in this respect that Type 13 (and 14) errors not only mean a substantial loss of information, they often prevent detection and/or correction of other errors.

A new NMC upper-air data decoder designed by Larry Sager is now undergoing testing, so that data holes at NMC should diminish once the new decoder is implemented.

3.3. "Canadian Fives".

Considering the variability of error frequency for each large region from month to month, one can see that this variability was, as a rule, comparatively small and rather irregular, and no improvement in the data quality during the year was evident. The only exception in that respect is the Block 71, Canada, where monthly numbers of hydrostatic errors were permanently decreasing, particularly during the first half of the year (see Fig. 10). This was connected with another finding in the course of the CHQC monitoring, even before it began to function operationally. A large number of comparatively small errors in the height of upper levels was detected, most often in height of the highest level (usually, 10 or 20 HPa), so that the DMA diagnosed them as Type 5 errors (that is why we called them "Canadian Fives").

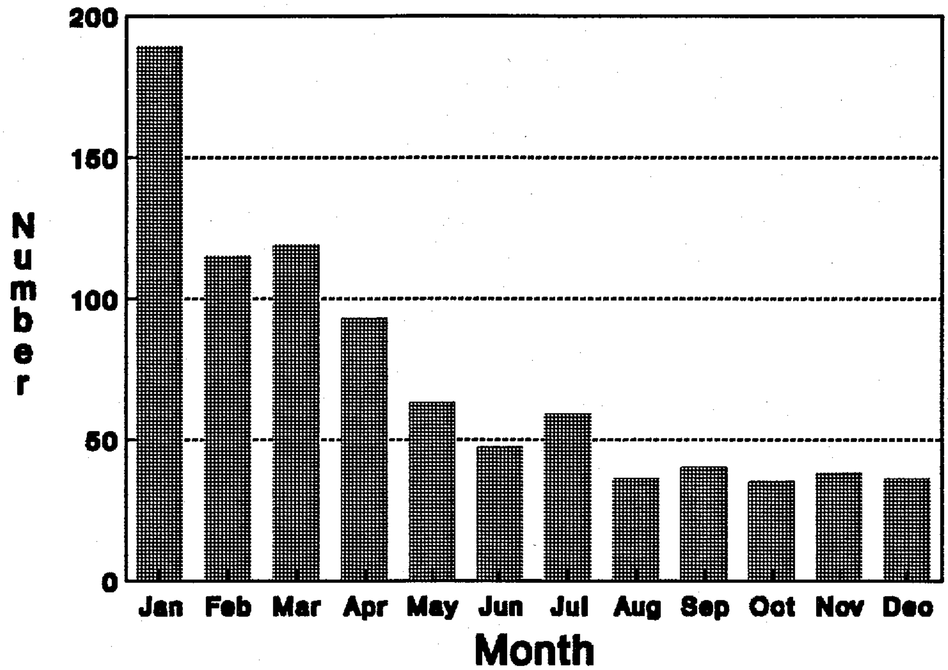

Fig.10. Monthly numbers of errors, Block 71.

All that looked rather strange, particularly because we knew that Canadian stations are equipped by microcomputers automatically processing rawinsonde measurements and also automatically putting the results into communication lines. The puzzle was, however, easily solved by our Canadian colleagues soon after we attracted their attention to it.

According to their processing code, if an observation level turns out to be very close to a mandatory one, namely, if its pressure differs by not more than 0.5 HPa from a mandatory pressure, then this level is just identified as the mandatory level. As long as balloons did not reach high elevations, this procedure did not cause any harm. However, for high elevations and, thus, low pressures, it may lead to substantial errors. It is easy to show, for example, that the resulting error in the 10 HPa height may exceed 200 m, and this is the order of magnitude of Type 5 errors over Canada detected by the CHQC. Once again, the Canadian processing code was implemented many years ago, and there is no doubt that errors of this type existed all the time since then, but they practically remained undetected before the CHQC began to operate.

Canadian specialists did everything possible in order to get rid of these errors. Unfortunately, their processing code is stored in the ROM (Read-Only Memory) of their microcomputers, and to modify it would be therefore too costly. The only remaining way was to perform some manual inspection in addition to the automatic processing, and that is what our Canadian colleagues did (and continue to do) quite efficiently.

This event has been a notable one in several respects. From the point of view of the DMA design, it demonstrated that the search of simple corrections is not universally applicable. From a purely technical point, it showed that one cannot exclude the probability that any operationally used code, however perfectly it seems to operate, could be modified at some future time. Most important, however, is that this event has demonstrated how productive the interaction between different scientific groups, even from different countries, can be. We are very grateful to our Canadian colleagues for their quick and professional reaction.

3.4. Permanent baseline errors.

Although the CHQC includes a baseline check, as described in CG, we decided not to use its results in the Decision Making Algorithm before other, statistical, checks are added to the hydrostatic check. The reason is that a baseline check residual may be due to several causes, and it is practically impossible to distinguish among them unless several checks, not only the hydrostatic one, are applied.

The baseline check results were, however, subjected to some quasi-operational monitoring. Some of the large baseline-check residuals were found to be caused by communication errors in heights of 1000 or 850 HPa surface, many others, we believe, are consequences of communication errors in surface air pressure, although we cannot be sure unless a horizontal check of the pressure is performed. We have also come across an event which may be seen, e.g., in Table 5 summarizing the baseline check results for a small sequence of observation times. Along with randomly occuring large residuals, there were some stations whose baseline check residuals were permanent and of approximately the same value. The only reasonable explanation of this phenomenon is that the data on station elevations (taken from the station dictionary) were wrong.

Attempts to examine, with the aid of some colleagues at NMC, what has actually happened to such stations, have shown that as many as three station dictionaries exist at NMC and their information does not always coincide! Detailed inspection of the dictionaries allowed our colleagues to change some of the elevations. For some other stations, however, no independent information has been found that the dictionary values for their elevation should, or may, be changed. These values have not yet been changed although, in our opinion, a permanent baseline error is, by itself, a sufficient reason to change the station elevation in all dictionaries.

Table 5 LARGE RESIDUALS OF THE BASELINE CHECK
June 21-26, 1989

	21/00	12	22/00	12	23/00	12	24/00	12	25/00	12	26/00
12374						-59					
13275	-43		-39		-37		-38				
16144	-43	-33	-37	-34	-40	-39	-42	-41	-35	-44	
16245				30							
16622		-37									
17030	-41	-42		-40	-41	-40		-40	-44	-43	
20107	51	50	53	49	54	54	55	50	54	57	54
20674									-47		
21358	-106	-100	-102	-102	-105	-101	-101	-104	-102	-101	-105
21965	34	34	40	31	39	34	37	32	34	33	39
22113	55	56	55	51							
23418	53	55	55	58	57	56	53	61	56	59	54
23921					-157						
24817					88						
24944					48						
27707	39	40	36	37	40	39	45	34	35	31	39
29838	-35		-36	-37	-36	-36	-35	-37		-36	-36
30309	80	86		83	82	83	86	90	89	86	82
36003									-31		
31329							-143				
31538	40		38	40	47	42	41	40		40	38
34152			-34								
35361								363			
36870	-186	-201	-182	-207	-179	-200	-181	-201	-187	-198	-182
37004							680				
37260	102	102	107	104	102	103	102	101	108	99	105
40080	-315										
40438		33									
41594											-197
41675			71								
41710						-660					
42182						-84					
42397								38			
42492							-50				
42971											-46
43014										61	
43128									-206		
43150		56							814		
43285										32	
43369									45		
48327				-90							
50953			-30								
51463		272	274	271	276	269	275	276	278	273	277
52203		64									
53513										95	
53845				105							
54161									634		
54337	-38	-39	-39	-39	-38		-32		-38		-38

Table 5, cont.

	21/00	12	22/00	12	23/00	12	24/00	12	25/00	12	26/00
54511							39	37	41	31	32
56294									-347		
57447	33	37	33			36	33	35			33
58666	-121	-118	-126		-115	-115	-117	-109	-122	-125	-124
58725	266			80							
59134	82	79	72	87	81	89	83	89	87	88	86
59316			51								
60571	44			39		30		35			
60630		-50									
60760		41		39		42		44			
61024		-486									
62053									-83		
71115						-88					
72291		-139	-146								
74671		75									
85799		-33		-32							
89050	36				37		32		40		34
89611						31					
91765		120		121							
94995	-42	-39			-41				-44	-36	-40
97072			73								
98327				-163	-165	-165					-160

Fortunately, the permanent baseline errors do not cause much harm, just because the data on station elevations are not used presently in any operational procedure except the baseline check. One cannot, however, exclude the possibility of such use in the future. In any case, it is necessary to have a standard dictionary of upper-air sounding stations, the same for all data assimilating centers over the world. The dictionary should be permanently updated reflecting operationally all changes in the station network.

3.5. The "Australian episode".

All problems described so far were permanent: each of them existed for a long time before they have been detected by the CHQC. In addition to such events, there were some other, non-permanent, events in 1989. We will describe only one, most notable, of such events, the one we called the Australian episode.

It occured suddenly, on the 26 of August. The number of Type 13 errors, the data holes, over the globe jumped from its usual values, about 5-10 per observation time, to a much higher level, and the numbers continued to be very high, particularly at 00 UTC, during subsequent days.

Analysis showed that this jump was caused by a dramatic increase of Type 13 error numbers over Australia (Block 94): being usually close to zero, these numbers grew to 13-20 at 00 UTC and 3-9 at 12 UTC (the difference is due to the fact that many Australian stations perform their rawinsonde observations once a day, at 00 UTC). Unlike usual data holes beginning somewhere whithin the part A of a report, each "Australian" hole began at the second reported level, so that all part A levels were missing except the lowest one. Moreover, there were reasons to believe that all Australian reports contained only one, lowest, level, and only a part of them was detected by the CHQC, because other reports simply did not have any part C from the beginning. As a result, the Australian reports were there, but each of them contained almost no information!

Thanks to Dr. Paul Julian, Chief of the NMC Quality Assurance Group, the cause of this event was found quite soon: everywhere in Australia, the coding of one digit, expressing the number of levels reporting winds, suddenly changed. Being thus unable to understand this digit, the NMC decoder ceased decoding of all other part A levels, as it always did (see section 3.2 above).

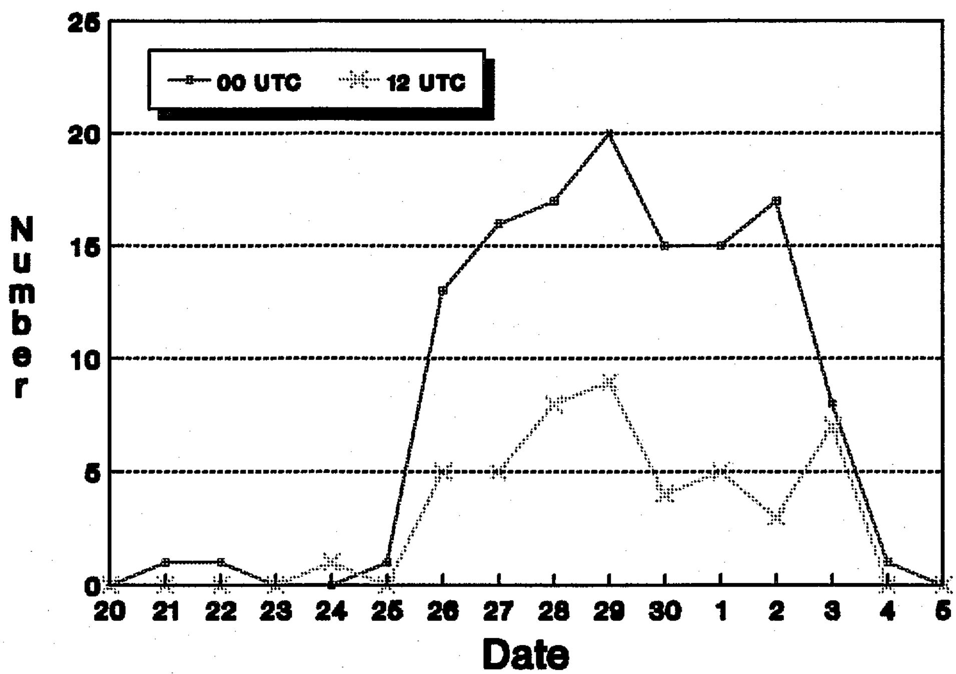

Fig.11. "Australian Episode"

Paul contacted Australian specialists, they returned to the proper coding, and the problem was fixed, so that the Australian episode lasted only for eight days (Fig. 11). If, however, there were no CHQC monitoring, no other routine means of data monitoring available at NMC would discover the episode, and it might have lasted much longer.

Several other situations occured during the year when the CHQC monitoring also detected some gaps in the data flow. These gaps, however, could be, and in some cases were, found by the routine NMC data quantity monitoring tools as well, and we shall not describe such situations.

4. Recommendations.

4.1. Applications of the CHQC code.

Several recommendations following from results of the Comprehensive Hydrostatic Quality Control monitoring have been already formulated above. Other recommendations, which are more general, will be described in this Section.

The first question we are going to consider is whether it is desirable to use the existing NMC CHQC code, as it is, at other centers involved in operational processing of rawinsonde data. We have to stress again in this respect that although the CHQC has been already used, quite successfully, at NMC for a comparatively long time, we are going to develop, as soon as possible, a multi-component Comprehensive Quality Control (CQC) of rawinsonde height and temperature, which will then replace the CHQC in NMC routine operations.

There exist, however, other centers, regional and local ones, where neither the numerical prediction nor objective analysis is being performed operationally, although they do receive and process rawinsonde reports. We believe that the CHQC code may be successfully applied at such centers after some minor changes or even without any change. The code is written in standard Fortran 77 and it is thoroughly documented, so that its implementation elsewhere will be comparatively easy. Even if a center already has some hydrostatic check code in operation, it seems desirable to replace it with the existing NMC CHQC code, just because of the high performance of the latter.

It is necessary to take into account, however, that as long as the Decision Making Algorithm is based only on the hydrostatic check residuals, it may sometimes, though very rarely, propose or even automatically perform wrong corrections. Remember that the CHQC, unlike any other existing QC procedure, never rejects any datum, it either corrects it or retains it as it was. Again, this is a consequence of the absence of other, statistical interpolation, checks in the CHQC: as soon as they are included, the DMA will be able to diagnose and automatically toss out incorrectable data. It will also deny corrections following from the the hydrostatic check residuals if they contradict the interpolation check results: such corrections will be never made.

The rare cases of wrong corrections happen, almost exclusively, when there are several suspected errors in one report or/and when the corrections are rather small, close to the margin between confident and non-confident corrections. For example, a detailed inspection of all Type 2 (temperature) corrections, equal to plus or minus 10° (amounting to 17% of the total number of temperature corrections), during November and December has shown that among 239 such corrections, 189 or 80% were definitely right, 25 (10%) were questionable, and 25 (10%) were wrong. This does not cause much harm, but still, one can avoid any wrong corrections by performing operational monitoring of the CHQC outputs, as is done by the NMC Meteorological Operations Division. This is rather easy to do because the numbers of proposed and introduced corrections are reasonable and because an overwhelming majority of automatically made corrections are quite obviously right.

4.2.Coding errors and code improvement.

Careful monitoring of the CHQC performance made it possible to recognize peculiar properties of some rough errors in rawinsonde reports. For example, we discovered that many one-digit corrections result in replacement of a digit by another digit looking like the reported one: 3 was often replaced by 5 or 8 (and vice versa), 7 was replaced by 2, and so on. This is easy to ascribe to what may be called transcription errors: a person wrote some digit in such a way that it has been understood as another digit. There were even cases when the same transcription error was made several times in a report or in a group of reports from the same block of stations.

Errors of another origin may be called coding errors: they were caused by wrong coding of messages. The most common coding error results in a wrong sign of temperature. As is well known, the sign is not reflected explicitly in coded messages, but instead is indicated by the last digit of temperature, expressing the tenths of degree: this digit should be even for a positive temperature and odd for a negative one. This artificial rule is often forgotten, and that is the main cause of sign errors.

There are many other more or less artificial rules in the existing code for upper-air sounding reports. For example, the first digit of a 700 HPa height expressing the number of kilometers is not included in coded messages and is to be added during their decoding, depending on the first digit (number of hectometers) in the transmitted part of the height. The same is true for some other heights. The arbitrarity of these rules has a double consequence: those who code messages often forget to omit the first digit for levels where this is necessary to do, and they often omit the first digit for levels where it is necessary not to do this. Unlike the temperature sign errors, it is not very easy to make precise corrections of these "under-omitting" and "over-omitting" errors.

Some other shortcomings of existing code were also identified as causes of information losses, as in the case of the "Australian episode" described above. Moreover, we came across some situations when both coding and decoding were performed perfectly, and still results were wrong. This happened at some Arctic and Antarctic stations in winter. If the 700 HPa height value at a station was, say, 2375 m, it was first, according to the coding rules, transmitted as 375 and then, following the decoding rules, made equal to 3375 m, i.e., higher by 1000 m. Of course, the CHQC DMA corrected all these "Arctic and Antarctic" errors.

There exist several ways to improve the present situation with coding errors. One can, for example, make the decoding algorithm latitude and/or season dependend in order to get rid of

Arctic and Antarctic errors. One can account for various kinds of coding errors explicitly in Decision Making Algorithms. The most desirable way is, however, to improve the existing code, particularly taking into account that it was designed several decades ago and has not undergone substantial changes since then.

At that time, the main requirement was to make the coded messages contain as few digits as possible. The reason for this was a severely limited capacity of communication channels. Many sacrifices were made to make the messages shorter, including omitting signs and some digits. As a consequence, the coding procedure is more complicated than it might be otherwise, and this is a source of many errors made in the course of manual coding.

The situation with communication lines has substantially improved during the last decades. Many channels are entirely automated or, better to say, computerized. They operate, of course, with binary digits, and there is no slightest need of any decimal coding for information following these channels. A special code, called the Binary Universal Form for the Representation of meteorological data (BUFR) has been designed to be applied under such conditions, and it is already used, quite successfully, for the information exchange between major prognostic centers. Undoubtedly, the BUFR code may, and should, be used for the transmission of data from stations to the centers, provided that all stages of the data flow operate entirely automatically, without any human intervention.

This state has been already achieved in some countries. Particularly, this is the case with all stations belonging to the National Weather Service of United States. However, the present situation is far from that in many countries covering large areas. Communication systems in such countries still require a great deal of manual operation, and this situation is not likely to essentially improve in the near future.

Having this in mind, it is natural to ask whether it is worthwhile, for the time being, to improve the existing code; to make it, so to say, less error-stimulating. Many specialists believe that there is no need to do so, and that the BUFR code may be successfully used for communicating the information from stations even if the communication process is not completely computerized, so that the BUFR data are transmitted and/or retransmitted manually.

We doubt this. Experience gained with analogous situations in some countries shows that such a "half-automation" can make things not better but worse. Under such circumstances, the numbers of communication related errors in BUFR data will be not lower but, likely, even higher, than they are now in much shorter, decimal messages. The only difference is that it will be more difficult to correct such errors.

We believe therefore that this problem deserves careful consideration. If, as seems most probable, the situation with communication lines in many countries is not expected to cardinally improve in coming years, then it is worthwhile to substantially improve the existing "manual" code thus making the probability of coding errors much lower than it is now.

The performance of an object or process usually depends on several factors, and one may try to improve the performance by modifying any of these factors. It usually happens, however, that one of the factors is "in a minimum", i.e., its present level is most insufficient for the object or process under consideration. Nothing essential may be achieved in such a situation by modifying other factors before the factor "in a minimum" is improved. To illustrate this principle, consider, for example, the observation network problem. If the density of an observation network is low, then one can achieve practically nothing by improving the accuracy of observations, it is necessary to create additional observations in order to make the network density sufficient. After this has been done, however, one cannot expect essential improvement by a further increase of the observation density, but in this situation, an increase in the accuracy may result in a substantial improvement of the network performance.

As another application of the principle one may take the performance of numerical weather prediction itself, with the important major factors of quality control, data assimilation, and forecast model. In the early days of NWP, and until recently, much attention was placed upon making the forecast model perform better. As forecast models have improved, the emphasis has been shared with the performance of the data assimilation system. This was necessary since the definition of the initial model state was the factor "in a minimum". The authors believe that the time will soon be here, or has already arrived, when the quality of the data is "in a minimum" for good forecasts, requiring increased quality control effort to become balanced with the other factors.

Regarding data quality, it is the communication network capacity that was, or at least was considered to be, in the minimum several decades ago. It would be wrong to say that the communication channel capacity, particularly the speed, is now sufficient everywhere, but our impression is that another factor has become the most limiting one nowadays, and this is the qualification of the personnel involved in processing and communication operations. The code improvement would substantially diminish the influence of this factor.

Another general problem that can also be mentioned in this respect is the problem of a choice between quick actions and thorough ones. It often happens that a case is found, more or less occasionally, when an existing system performs insufficiently. Reacting to such a case, one may quickly modify some parameters of the system in the hope that the modified system will perform well in this case as well as in other, analogous, cases. A thorough investigation may be undertaken instead resulting in more reasonable modification of the system but requiring much time.

Analysis of modifications performed with operational data assimilation and numerical prediction systems shows that many such modifications are made too quickly, just in order to get rid of a poor behaviour of the system as soon as possible. This is, of course, easy to understand, but the price to be paid for such a hasty decision is often too high: other shortcomings become apparent taking place despite the modification, or even due to it, so that the problem becomes even more complicated than it was before the modification.

Situations of an opposite character also happen sometimes, when a necessary action is delayed for a long time or not undertaken at all because some other actions in a close or distant future will automatically solve the problem. The situation with "data holes" described above (Section 3. 2) is a good example of this kind. Our impression based on the CHQC monitoring is that the present situation with the upper-air sounding code belongs to the same category.

It is necessary, however, to warn against too quick decisions, those resulting in only some minor changes in the existing coding system. The leading principle is to be changed: instead of minimizing the number of digits, the improved coding system should maximize the clarity of the coding procedure and, thus, the clarity of coded messages.

4.3. Desirability of independent measurements of height, pressure, and temperature.

The last proposal we feel desirable to consider here deals not so much with any kind of the observational data processing as with the observations themselves.

The hydrostatic equation connects three meteorological parameters: pressure, height, and (virtual) temperature. All three are measurable in the course of rawinsounding, and all are actually being measured in a majority of countries. Radar measurements of the sonde positions are performed in order to compute wind speed and direction, but the sonde heights, easily obtainable in the course of these measurements, are not used to compute heights of mandatory (or any other) surfaces. The hydrostatic equation is used instead in order to compute the heights.

This situation may be explained from a historical point of view. For a long time, only angular coordinates of balloons were measured, first by ordinary theodolites, then by radio-theodolites. Distance measurements by radars allowing the determination of heights were introduced later, and the accuracy of these, direct, height observations was inferior to that of the indirect method to compute heights by the hydrostatic equation. Many specialists believe that this is still the case.

We don't know whether this is true. What we do know is that Russian specialists came to an opposite conclusion almost a decade ago and decided to use radar-measured heights directly. Unfortunately, the decision was to use these heights not additionally, but instead of measuring pressure. Russian rawinsondes simply have no pressure sensors. The hydrostatic equation is thus used twice: first to compute pressure for each observation level, and then to compute heights of mandatory isobaric surfaces by the usual procedure of accumulating thicknesses from the lowest level upwards.

The accumulation procedure, common to all rawinsonde systems, has, however, many undesirable consequences. First of all, it results in a very strong vertical correlation of the height "observation" (better to say, indirect computation) errors. It is easy to show that, even if the temperature observation errors are vertically uncorrelated, the accumulation procedure leads, for example, to correlation coefficients between 100 and 50 HPa height errors exceeding 80%. This means that the information contained in a set of mandatory level heights over a station exceeds only slightly the information contained in one or two such heights. As a result of this, the optimum interpolation in the course of the data assimilation, being formally three-dimensional, turns out practically to be two-dimensional almost everywhere as long as isobaric height is concerned.

The second consequence of using hydrostatically computed heights instead of measured ones is the fact that even small errors in temperature may lead to large errors in heights if the temperature errors are vertically correlated. For example, a systematic error (a "bias") in temperature as small as 0.5^{o} results in the 30 HPa height error exceeding 50 m. This undesirable effect is particularly pronounced near the boundaries between areas using different rawinsonde types. Their temperature sensors usually react slightly different to solar radiation, and such small but persistent temperature differences often result in large fictitious horizontal gradients in (computed) heights, particularly in the stratosphere. According to Paul Julian (personal communication), these fictitious gradients are often so high, that they may even prevent successful detection of rough errors by means of a horizontal check.

The third consequence of using computed heights was already mentioned above: as long as this is the case, the CHQC is unable to detect observation errors, it simply does not react to them.

If all three parameters were taken from observations, then none of the listed effects would take place or, at least, they would be much less pronounced. It is necessary to take into account, however, that height, pressure and virtual temperature will be hydrostatically uncoupled due to random observational errors, and they have therefore to be, first of all, hydrostatically adjusted to each other. The necessity to do so may seem not evident, particularly if we remember that the hydrostatic equation should not be obeyed with absolute accuracy because this equation is just the rather simplified form of the vertical projection of the momentum equation, resulting from neglecting some its terms, including the vertical acceleration. Detailed analysis shows, however, that neglected terms of this equation are small in comparison not only with its main terms, but also with its residuals caused by random errors in observed data. This justifies the necessity of the hydrostatic adjustment.

We are not going to consider here the procedure of this adjustment. Several points, however, deserve to be mentioned. The adjustment procedure is simple and requires very little computer time. It results, first of all, in a marked increase in accuracy of all three parameters. It does not produce any substantial vertical correlation of random errors in adjusted values. And, finally, the hydrostatic adjustment is, by itself, a good instrument to detect rough observational errors. The leading principle of doing so is simple: if all hydrostatic residuals are reasonably small, use them to hydrostatically adjust the data; if, however, one of the residuals is, or several are, unreasonably large, then search for rough observational errors in the data.

5. Summary

The Comprehensive Hydrostatic Quality Control proves to be an effective means for detecting rough errors in rawinsonde reports and correcting many of them, as well as for the data quality monitoring. The CHQC productivity is due, most of all, to its advanced Decision Making Algorithm described in detail in CG. The main distinctive property of this DMA is that it attempts to determine the cause of each suspected error and to use this information in order to decide how, and if, to correct it.

Several improvements of the CHQC, and particularly of its DMA have been made in the course of its operational use, and we believe that no further improvements are needed as long as all DMA actions are based on the hydrostatic check alone.

At the same time, the CHQC monitoring results show quite definitely that the performance of a Complex Quality Control (CQC) containing statistical interpolation checks along with the hydrostatic one will be substantially higher than that of the CHQC. The number of cases, when the DMA is incapable of making a univalued decision and human intervention is needed, will be much smaller. This is particularly important, because experience shows that it is very difficult for NMC Meteorological Operation Division specialists, even when they have the CHQC outputs, to perform a thorough manual analysis of questionable cases - just because of severe time limitations.

It is perhaps even more important that the CQC will be able to detect rough observation errors. We know very little as yet about the frequency of such errors, because the CHQC does not detect them. However, we came across several cases when a report was distorted by communication error(s) (and was therefore caught by the CHQC) and by evidently existing measurement errors. This is an indication, though a weak one, that rough errors of observational origin occur not very seldom.

As mentioned before, the CQC design is now underway, and we hope to be able in several months to begin testing it.

Acknowledgments. In the course of our CHQC monitoring, we contacted with many colleagues, both at NMC and elsewhere, and we are grateful to all of them. We wish to express our particular gratitude to Dr. William Bonner for his encouragement and support, to Dr. Paul Julian for his constructive comments, and to Lauren Morone whose thorough editorial work resulted in marked improvement of the text.

References.

Bottger, H., A. Radford, and D. Soderman, 1987: ECMWF monitoring tools and their application to North American radiosonde data. European Centre for Medium Range Weather Forecasts, Operations Department, Technical Memorandum No 133.

Collins W. G. and L. S. Gandin, 1990: Comprehensive Hydrostatic Quality Control at the National Meteorological Center. (Submitted for publication in Monthly Weather Review.)

Hollingsworth A., D. B. Shaw, P. Lonnberg, L. Illari, K. Arpe and A. J. Simmons, 1986: Monitoring of observation and analysis quality by a data assimilation system. Mon. Wea. Rev., **114**, 862-879.

Appendix: Tables of monthly statistics.

The tables on the following pages contain more detailed statistics for possible use by specialists directly involved in the design and/or monitoring of the quality control algorithms. Each table contains monthly averaged, or summed up, statistics grouped either by error types or by large regions. Values averaged over the whole year (denoted "yr avg") are also presented, together with the standard deviations of monthly values ("yr std"). A separate table contains monthly numbers of rawinsonde reports received at NMC.

Some of statistics were not collected for the first 4 or 6 months of the year, and corresponding data are missing in the tables.

MONTHLY ABSOLUTE NUMBERS OF ERRORS PER OBSERVATION TIME BY ERROR TYPE

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	yr avg	yr std
1	1010	883	903	788	803	803	901	735	876	838	901	912	863	70.1
2	737	674	725	638	604	648	791	727	729	711	776	662	702	54.5
1+2	1747	1557	1628	1426	1407	1451	1692	1462	1605	1549	1677	1574	1565	106.3
Other	2176	1747	1620	1334	1190	1321	1622	1175	1517	1378	1214	1102	1450	293.5
3	297	292	258	224	236	197	100	76	100	101	54	65	167	88.7
4	393	307	257	226	200	228	132	70	131	103	80	103	186	96.0
5	506	298	247	223	226	235	180	121	160	167	130	116	217	102.3
4+5	899	605	504	449	426	463	312	191	291	270	210	219	403	195.3
6	477	400	399	305	231	323	209	175	182	160	167	103	261	112.6
7-10	134	130	70	80	75	96	126	93	119	100	114	99	103	20.8
11	252	221	240	182	169	175	228	152	185	180	170	190	195	30.4
13							408	343	470	418	343	310	382	54.7
14							40	29	38	37	38	35	36	3.5
22							42	44	50	36	51	27	42	8.3

MONTHY ABSOLUTE NUMBERS OF ERRORS

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	yr avg	yr std
Overall	3923	3304	3248	2760	2597	2772	3314	2637	3122	2927	2891	2676	3014	370.0
W Europe	82	85	64	64	40	99	77	62	84	68	68	59	71	14.7
E Europe	140	102	134	78	79	99	140	103	143	90	102	93	109	23.1
USSR	773	630	695	484	510	520	726	526	647	589	605	595	608	86.9
W Asia	103	84	75	105	66	69	85	93	67	81	94	72	83	13.0
Indo Ch	499	516	467	486	407	424	472	436	471	489	438	246	446	67.7
Mongolia	126	109	98	84	98	107	98	94	100	107	99	117	103	10.5
Japan	85	86	78	76	78	66	129	93	96	99	109	116	93	17.6
Indonesia	146	160	121	118	73	78	134	90	96	97	93	132	112	26.6
China	741	572	586	439	447	548	632	501	502	494	500	532	541	80.4
Africa	236	187	186	192	168	185	157	158	221	183	204	228	192	24.8
S Africa	16	14	3	2	7	2	5	9	9	18	7	11	9	5.1
USA	41	52	25	23	17	26	24	19	23	30	28	18	27	9.7
Canada	189	115	119	93	63	47	59	36	40	35	38	36	73	45.8
C America	140	70	104	79	84	91	71	46	49	47	57	62	75	26.2
S America	185	181	143	177	146	125	127	107	145	124	98	96	138	29.6
Antarctic	19	32	24	20	31	16	24	17	17	33	19	12	22	6.6
Pac Isl	224	185	165	120	161	159	196	133	197	210	210	140	175	32.2
Australia	178	124	161	120	122	111	158	114	215	133	122	111	139	31.0

MONTHLY ABSOLUTE NUMBERS OF RECEIVED REPORTS

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	yr avg	yr std
Land reports					44535	41991	43251	37676	42393	43093	42686	43475	42387.5	1919.7
Ship reports					919	964	845	806	808	972	819	705	854.8	85.3

MEAN NO. OF ERRORS PER OBSERVATION TIME BY ERROR TYPE

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	yr avg	yr std
1	16.3	15.8	15.1	13.6	13.2	13.4	14.5	12.9	14.6	13.7	15.0	14.7	14.4	1.0
2	11.9	12.0	12.1	11.0	9.9	10.8	12.8	12.8	12.2	11.7	12.9	10.7	11.7	0.9
1+2	28.2	27.8	27.1	24.6	23.1	24.2	27.3	25.6	26.8	25.4	28.0	25.4	26.1	1.6
Other	35.1	31.2	27.0	23.0	19.5	22.0	26.2	20.6	25.3	22.6	20.2	17.8	24.2	4.8
3	4.8	5.2	4.3	3.9	3.9	3.3	1.6	1.3	1.7	1.7	0.9	1.0	2.8	1.5
4	6.3	5.5	4.3	3.9	3.3	3.8	2.1	1.2	2.2	1.7	1.3	1.7	3.1	1.6
5	8.2	5.3	4.1	3.8	3.7	3.9	2.9	2.1	2.7	2.7	2.2	1.9	3.6	1.7
4+5	14.5	10.8	8.4	7.7	7.0	7.7	5.0	3.4	4.9	4.4	3.5	3.5	6.7	3.2
6	7.7	7.1	6.7	5.3	3.8	5.4	3.4	3.1	3.0	2.6	2.8	1.7	4.4	1.9
7-10	2.2	2.3	1.2	1.4	1.2	1.6	2.0	1.6	2.0	1.6	1.9	1.6	1.7	0.3
11	4.1	3.9	4.0	3.1	2.8	2.9	3.7	2.7	3.1	3.0	2.8	3.1	3.3	0.5
13							6.6	6.0	7.8	6.9	5.7	5.0	6.3	0.9
14							0.6	0.5	0.6	0.6	0.6	0.6	0.6	0.0
22							0.7	0.8	0.8	0.6	0.9	0.4	0.7	0.1

MEAN NO. OF ERRORS PER OBSERVATION TIME BY LARGE REGIONS

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	yr avg	yr std
Overall	63.3	59.0	54.1	47.6	42.6	46.2	53.5	46.3	52.0	48.0	48.2	43.2	50.3	6.0
W Europe	1.3	1.5	1.1	1.1	0.7	1.7	1.2	1.1	1.4	1.1	1.1	1.0	1.2	0.3
E Europe	2.3	1.8	2.2	1.3	1.3	1.7	2.3	1.8	2.4	1.5	1.7	1.5	1.8	0.4
USSR	12.5	11.3	11.6	8.3	8.4	8.7	11.7	9.2	10.8	9.7	10.1	9.6	10.1	1.3
W Asia	1.7	1.5	1.3	1.8	1.1	1.2	1.4	1.6	1.1	1.3	1.6	1.2	1.4	0.2
Indo Ch	8.0	9.2	7.8	8.4	6.7	7.1	7.6	7.6	7.9	8.0	7.3	4.0	7.5	1.2
Mongolia	2.0	1.9	1.6	1.4	1.6	1.8	1.6	1.6	1.7	1.8	1.7	1.9	1.7	0.2
Japan	1.4	1.5	1.3	1.3	1.3	1.1	2.1	1.6	1.6	1.6	1.8	1.9	1.5	0.3
Indonesia	2.4	2.9	2.0	2.0	1.2	1.3	2.2	1.6	1.6	1.6	1.6	2.1	1.9	0.5
China	12.0	10.2	9.8	7.6	7.3	9.1	10.2	8.8	8.4	8.1	8.3	8.6	9.0	1.3
Africa	3.8	3.3	3.1	3.3	2.8	3.1	2.5	2.8	3.7	3.0	3.4	3.7	3.2	0.4
S Africa	0.3	0.3	0.1	0.0	0.1	0.0	0.1	0.2	0.2	0.3	0.1	0.2	0.1	0.1
USA	0.7	0.9	0.4	0.4	0.3	0.4	0.4	0.3	0.4	0.5	0.5	0.3	0.5	0.2
Canada	3.0	2.1	2.0	1.6	1.0	0.8	1.0	0.6	0.7	0.6	0.6	0.6	1.2	0.8
C America	2.3	1.3	1.7	1.4	1.4	1.5	1.1	0.8	0.8	0.8	1.0	1.0	1.2	0.4
S America	3.0	3.2	2.4	3.1	2.4	2.1	2.0	1.9	2.4	2.0	1.6	1.5	2.3	0.5
Antarctic	0.3	0.6	0.4	0.3	0.5	0.3	0.4	0.3	0.3	0.5	0.3	0.2	0.4	0.1
Pac Isl	3.6	3.3	2.8	2.1	2.6	2.7	3.2	2.3	3.3	3.4	3.5	2.3	2.9	0.5
Australia	2.9	2.2	2.7	2.1	2.0	1.9	2.5	2.0	3.6	2.2	2.0	1.8	2.3	0.5

PERCENT OF ERRORS BY TYPES

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	yr avg	yr std
1	25.7	26.7	27.8	28.6	30.9	29.0	27.2	27.9	28.1	28.6	31.2	34.1	28.8	2.2
2	18.8	20.4	22.3	23.1	23.3	23.4	23.9	27.6	23.4	24.3	26.8	24.7	23.5	2.3
1+2	44.5	47.1	50.1	51.7	54.2	52.3	51.1	55.4	51.4	52.9	58.0	58.8	52.3	3.9
Other	55.5	52.9	49.9	48.3	45.8	47.7	48.9	44.6	48.6	47.1	42.0	41.2	47.7	3.9
3	7.6	8.8	7.9	8.1	9.1	7.1	3.0	2.9	3.2	3.5	1.9	2.4	5.5	2.7
4	10.0	9.3	7.9	8.2	7.7	8.2	4.0	2.7	4.2	3.5	2.8	3.8	6.0	2.6
5	12.9	9.0	7.6	8.1	8.7	8.5	5.4	4.6	5.1	5.7	4.5	4.3	7.0	2.5
4+5	22.9	18.3	15.5	16.3	16.4	16.7	9.4	7.2	9.3	9.2	7.3	8.2	13.1	5.0
6	12.2	12.1	12.3	11.1	8.9	11.7	6.3	6.6	5.8	5.5	5.8	3.8	8.5	3.0
7-10	3.4	3.9	2.2	2.9	2.9	3.5	3.8	3.5	3.8	3.4	3.9	3.7	3.4	0.5
11	6.4	6.7	7.4	6.6	6.5	6.3	6.9	5.8	5.9	6.1	5.9	7.1	6.5	0.5
13							12.3	13.0	15.1	14.3	11.9	11.6	13.0	1.3
14							1.2	1.1	1.2	1.3	1.3	1.3	1.2	0.1
22							1.3	1.7	1.6	1.2	1.8	1.0	1.4	0.3

PERCENT OF ERRORS BY LARGE REGIONS

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	yr avg	yr std
W Europe	2.1	2.6	2.0	2.3	1.5	3.6	2.3	2.4	2.7	2.3	2.4	2.2	2.4	0.5
E Europe	3.6	3.1	4.1	2.8	3.0	3.6	4.2	3.9	4.6	3.1	3.5	3.5	3.6	0.5
USSR	19.7	19.1	21.4	17.5	19.6	18.8	21.9	19.9	20.7	20.1	20.9	22.2	20.2	1.3
W Asia	2.6	2.5	2.3	3.8	2.5	2.5	2.6	3.5	2.1	2.8	3.3	2.7	2.8	0.5
Indo Ch	12.7	15.6	14.4	17.6	15.7	15.3	14.2	16.5	15.1	16.7	15.2	9.2	14.9	2.1
Mongolia	3.2	3.3	3.0	3.0	3.8	3.9	3.0	3.6	3.2	3.7	3.4	4.4	3.4	0.4
Japan	2.2	2.6	2.4	2.8	3.0	2.4	3.9	3.5	3.1	3.4	3.8	4.3	3.1	0.7
Indonesia	3.7	4.8	3.7	4.3	2.8	2.8	4.0	3.4	3.1	3.3	3.2	4.9	3.7	0.7
China	18.9	17.3	18.0	15.9	17.2	19.8	19.1	19.0	16.1	16.9	17.3	19.9	17.9	1.3
Africa	6.0	5.7	5.7	7.0	6.5	6.7	4.7	6.0	7.1	6.3	7.1	8.5	6.4	0.9
S Africa	0.4	0.4	0.1	0.1	0.3	0.1	0.2	0.3	0.3	0.6	0.2	0.4	0.3	0.2
USA	1.0	1.6	0.8	0.8	0.7	0.9	0.7	0.7	0.7	1.0	1.0	0.7	0.9	0.2
Canada	4.8	3.5	3.7	3.4	2.4	1.7	1.8	1.4	1.3	1.2	1.3	1.3	2.3	1.2
C America	3.6	2.1	3.2	2.9	3.2	3.3	2.1	1.7	1.6	1.6	2.0	2.3	2.5	0.7
S America	4.7	5.5	4.4	6.4	5.6	4.5	3.8	4.1	4.6	4.2	3.4	3.6	4.6	0.8
Antarctic	0.5	1.0	0.7	0.7	1.2	0.6	0.7	0.6	0.5	1.1	0.7	0.4	0.7	0.2
Pac Isl	5.7	5.6	5.1	4.3	6.2	5.7	5.9	5.0	6.3	7.2	7.3	5.2	5.8	0.8
Australia	4.5	3.8	5.0	4.3	4.7	4.0	4.8	4.3	6.9	4.5	4.2	4.1	4.6	0.8

PERCENT OF CONFIDENTLY CORRECTED ERRORS BY LARGE REGIONS

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	yr avg	yr std
Overall	45	46	50	51	53	52	55	59	55	56	62	63	53.9	5.3
W Europe	59	66	67	56	68	57	68	66	74	60	63	70	64.5	5.3
E Europe	53	63	0	63	63	64	60	56	64	60	72	70	57.3	18.0
USSR	53	55	8	62	61	59	61	67	61	66	72	68	57.8	15.9
W Asia	43	43	41	52	48	42	38	43	54	53	49	57	46.9	5.8
Indo Ch	29	31	31	32	40	36	39	42	38	35	39	33	35.4	4.1
Mongolia	44	50	57	52	49	56	62	60	68	64	56	69	57.3	7.4
Japan	73	56	56	61	64	62	50	58	56	72	64	65	61.4	6.5
Indonesia	35	42	36	35	58	44	39	54	44	50	59	48	45.3	8.2
China	49	49	57	59	60	58	63	69	61	64	74	69	61.0	7.3
Africa	42	50	41	43	44	48	46	54	57	46	56	54	48.4	5.4
S Africa	38	21	33	0	29	50	20	22	67	28	43	55	33.8	17.3
USA	49	19	30	43	65	46	46	37	44	40	46	39	42.0	10.6
Canada	32	32	45	43	38	57	46	47	60	46	32	58	44.7	9.6
C America	42	59	62	42	54	53	61	54	65	51	58	58	54.9	6.9
S America	37	43	50	58	52	53	51	62	57	57	56	49	52.1	6.6
Antarctic	47	34	38	80	74	50	75	65	88	61	37	58	58.9	17.3
Pac Isl	33	43	47	46	45	35	42	45	48	54	50	57	45.4	6.6
Australia	52	53	58	70	61	63	61	58	32	45	72	73	58.2	11.2

PERCENT OF ERROR-CONTAINING REPORTS BY LARGE REGIONS

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	yr avg	yr std
Overall					5.8	6.6	7.7	7.0	7.4	6.8	6.8	6.2	6.8	0.6
W Europe	2	2	2	2	1	2	2	2	2	2	2	1	1.8	0.4
E Europe	9	8	9	6	5	7	10	8	10	6	7	6	7.6	1.6
USSR	7	7	7	4	5	5	7	6	6	6	6	6	6.0	0.9
W Asia	9	8	7	10	6	7	9	11	7	8	10	7	8.3	1.5
Indo Ch	31	35	29	32	27	29	29	31	30	29	29	25	29.7	2.4
Mongolia	25	24	20	18	20	24	21	25	23	22	23	24	22.4	2.1
Japan	5	5	4	4	4	4	7	6	5	5	6	6	5.1	1.0
Indonesia	29	33	25	25	16	15	22	21	17	16	16	23	21.5	5.5
China	10	9	8	6	6	8	8	8	7	7	7	7	7.6	1.1
Africa	20	17	16	17	14	18	14	15	18	16	17	17	16.6	1.7
S Africa	2	2	0.4	0.3	1	0.3	0.7	1.4	1.3	2.5	1	1.5	1.2	0.7
USA	1	1	0.4	0.4	0.3	0.5	0.4	0.4	0.4	0.6	0.5	0.3	0.5	0.2
Canada	10	6	6	5	3	2	3	2	2	2	2	2	3.8	2.4
C America	15	8	11	9	8	9	8	6	7	6	7	8	8.5	2.4
S America	19	20	15	19	14	15	16	14	17	16	14	12	15.9	2.3
Antarctic	3	6	4	4	5	3	4	3	4	6	3	2	3.9	1.2
Pac Isl	15	14	11	8	13	12	15	13	15	14	14	10	12.8	2.1
Australia	12	9	11	8	8	7	10	9	14	9	8	7	9.3	2.0